

500P15710300

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

JC654 U.S. PTO  
09/749345  
12/27/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年12月28日

出 願 番 号

Application Number:

平成11年特許願第373780号

出 願 人

Applicant(s):

ソニー株式会社



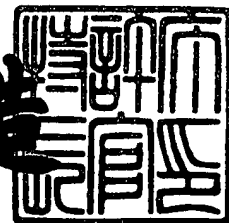
CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年10月13日



特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2000-3085021

【書類名】 特許願

【整理番号】 9900958303

【提出日】 平成11年12月28日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 5/02

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 島川 真人

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 山崎 信英

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 小林 恵理香

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 赤羽 誠

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 小林 賢一郎

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 山田 敬一

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 新田 朋晃

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代表者】 出井 伸之

【代理人】

【識別番号】 100082131

【弁理士】

【氏名又は名称】 稲本 義雄

【電話番号】 03-3369-6479

【手数料の表示】

【予納台帳番号】 032089

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9708842

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成装置および方法、並びに記録媒体

【特許請求の範囲】

【請求項 1】 対象文に対応して音声信号を合成する音声合成装置において

動作モデルに従って動作状態を遷移する動作状態遷移手段と、

感情モデルに従って感情状態を遷移する感情状態遷移手段と、

前記動作状態および前記感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択手段と、

前記選択手段が選択した前記制御情報に含まれる音声合成情報に基づき、前記対象文に対応して音声信号を合成する合成手段と

を含むことを特徴とする音声合成装置。

【請求項 2】 前記音声合成情報には、素片データ ID、音節セット ID、声の高さパラメータ、アクセントの強さパラメータ、声立ての強さパラメータ、および発話スピードパラメータのうち、少なくとも 1 つが含まれる

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 3】 外部の状況を検出する検出手段をさらに含み、

前記選択手段は、前記検出手段の検出結果にも基づいて前記制御情報を選択する

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 4】 個体情報を保持する保持手段をさらに含み、

前記選択手段は、前記保持手段が保持する前記個体情報にも基づいて前記制御情報を選択する

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 5】 起動以降の経過時間をカウントするカウント手段をさらに含み、

前記選択手段は、前記カウント手段がカウントした前記経過時間にも基づいて前記制御情報を選択する

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 6】 前記動作状態遷移手段が前記動作状態を遷移した回数、および前記感情状態遷移手段が前記感情状態を遷移した回数のうち、少なくとも一方を累積する累積手段をさらに含み、

前記選択手段は、前記累積手段が累積した回数にも基づいて前記制御情報を選択する

ことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 7】 前記選択手段が選択した前記制御情報に含まれる選択情報に対応する単語置換辞書を用いて、前記対象文に含まれる単語を置換する置換手段を

さらに含むことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 8】 前記選択手段が選択した前記制御情報に含まれる選択情報に対応する文体変換規則に従って、前記対象文の文体を変換する変換手段を

さらに含むことを特徴とする請求項 1 に記載の音声合成装置。

【請求項 9】 前記音声合成装置はロボットであることを特徴とする請求項 1 に記載の音声合成装置。

【請求項 10】 対象文に対応して音声信号を合成する音声合成装置の音声合成方法において、

動作モデルに従って動作状態を遷移する動作状態遷移ステップと、

感情モデルに従って感情状態を遷移する感情状態遷移ステップと、

前記動作状態および前記感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、

前記選択ステップの処理で選択された前記制御情報に含まれる音声合成情報に基づき、前記対象文に対応して音声信号を合成する合成ステップと

を含むことを特徴とする音声合成方法。

【請求項 11】 対象文に対応して音声信号を合成する音声合成用のプログラムであって、

動作モデルに従って動作状態を遷移する動作状態遷移ステップと、

感情モデルに従って感情状態を遷移する感情状態遷移ステップと、

前記動作状態および前記感情状態のうち、少なくとも一方に基づいて制御情報

を選択する選択ステップと、

前記選択ステップの処理で選択された前記制御情報に含まれる音声合成情報に基づき、前記対象文に対応して音声信号を合成する合成ステップと

を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声合成装置および方法、並びに記録媒体に関し、例えば、ロボットに搭載して、当該ロボットの感情や動作等に対応して合成する音声信号を変化させる場合に用いて好適な音声合成装置および方法、並びに記録媒体に関する。

【0002】

【従来の技術】

従来、言葉を発声するロボットは存在する。例えば、言葉を発声するロボットが、自己の感情を変化させて、感情に応じて話し方を変えるようになり、例えばまた、ロボットに、種別、性別、年齢、出身地、性格、体格などの個性を設定し、個性に対応して話し方を変化させるようになれば、よりリアルに生命体を模擬したものとなる。

【0003】

【発明が解決しようとする課題】

そのようなロボットに対してユーザは、あたかも生きたペットに対するような親しみや愛情をもって接することができるようになるが、そのようなロボットは実現されていない課題があった。

【0004】

本発明はこのような状況に鑑みてなされたものであり、感情や動作等に対応して話し方を変化させるようにすることにより、よりリアルに生命体を模擬したロボットを実現することを目的とする。

【0005】

【課題を解決するための手段】

本発明の音声合成装置は、動作モデルに従って動作状態を遷移する動作状態遷移手段と、感情モデルに従って感情状態を遷移する感情状態遷移手段と、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択手段と、選択手段が選択した制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成する合成手段とを含むことを特徴とする。

## 【0006】

本発明の音声合成装置は、外部の状況を検出する検出手段をさらに含むことができ、前記選択手段は、検出手段の検出結果にも基づいて制御情報を選択するようにすることができる。

## 【0007】

本発明の音声合成装置は、個体情報を保持する保持手段をさらに含むことができ、前記選択手段は、保持手段が保持する個体情報にも基づいて制御情報を選択するようにすることができる。

## 【0008】

本発明の音声合成装置は、起動以降の経過時間をカウントするカウント手段をさらに含むことができ、前記選択手段は、カウント手段がカウントした経過時間にも基づいて制御情報を選択するようにすることができる。

## 【0009】

本発明の音声合成装置は、動作状態遷移手段が動作状態を遷移した回数、および感情状態遷移手段が感情状態を遷移した回数のうち、少なくとも一方を累積する累積手段をさらに含むことができ、前記選択手段は、累積手段が累積した回数にも基づいて制御情報を選択するようにすることができる。

## 【0010】

本発明の音声合成装置は、選択手段が選択した制御情報に含まれる選択情報に対応する単語置換辞書を用いて、対象文に含まれる単語を置換する置換手段をさらに含むことができる。

## 【0011】

本発明の音声合成装置は、選択手段が選択した制御情報に含まれる選択情報に対応する文体変換規則に従って、対象文の文体を変換する変換手段をさらに含む

ことができる。

【0012】

本発明の音声合成方法は、動作モデルに従って動作状態を遷移する動作状態遷移ステップと、感情モデルに従って感情状態を遷移する感情状態遷移ステップと、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、選択ステップの処理で選択された制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成する合成ステップとを含むことを特徴とする。

【0013】

本発明の記録媒体のプログラムは、動作モデルに従って動作状態を遷移する動作状態遷移ステップと、感情モデルに従って感情状態を遷移する感情状態遷移ステップと、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、選択ステップの処理で選択された制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成する合成ステップとを含むことを特徴とする。

【0014】

本発明の音声合成装置、音声合成方法、および記録媒体のプログラムにおいては、動作モデルに従って動作状態が遷移され、感情モデルに従って感情状態が遷移される。また、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報が選択され、選択された制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号が合成される。

【0015】

【発明の実施の形態】

図1は、本発明を適用したロボットの音声合成に関わる部分の構成例を示している。このロボットは、発話機能を有し、かつ、感情と動作を変化させて、感情や動作の変化に応じて話し方を変化するようになされている。

【0016】

各種センサ1は、ロボット外部の状況やロボットに対する作用を検出して、検出結果をロボット運動系制御部10に出力する。例えば、外気温センサ2は、ロ



ボットの外気温を検出する。温度センサ 3 と接触センサ 4 は、対をなして近傍に設けられており、接触センサ 4 がロボットに対する物体の接触を検知し、温度センサ 3 が接触した物体の温度を検出する。感圧センサ 5 は、ロボットに加えられる外力（例えば、叩かれる力や撫でられる力など）の強さを検出する。風速センサ 6 は、ロボット外部の風速を検出する。照度センサ 7 は、ロボット外部の照度を検出する。画像センサ 8 は、例えば CCD などからなり、ロボット外部の景色を画像信号として検出する。音声センサ 9 は、例えばマイクロフォンからなり、音声を検出する。

## 【0017】

ロボット運動系制御部 10 は、図 2 に示すように、運動系処理部 31 と動作モデル 32 から構成されており、当該ロボットの動作を管理する。運動系処理部 31 は、各種センサ 1 から入力される検出結果、自己の内部で生成する内部イベント、および、ロボット思考系制御部 11 から入力される指令を動作モデル 32 に照らし合わせてロボットの動作を遷移させ、現在の動作の状態を動作ステートとして音声合成コントロール情報選択部 12 に出力する。運動系処理部 31 はまた、例えば、感圧センサ 5 の検知結果が所定の閾値以上の力であったとき、動作イベントとして「頭を叩かれた」と判定するように、各種センサ 1 から入力される検出結果に基づき、動作イベントを判定してロボット思考系制御部 11 に出力する。運動系処理部 31 はさらに、各種センサ 1 からの検出結果をロボット思考系制御部 11 に中継する。なお、各種センサ 1 から思考系処理部 41 に検出結果を直接入力するようにしてもよい。

## 【0018】

動作モデル 32 には、図 3 に示すように、標準状態から各種の動作に遷移するときの条件が記述されている。例えば、標準状態において、指令「歩く」が命令された場合、動作「歩く」に遷移することや、例えばまた、指令「起き上がる」が命令された場合、動作「起き上がる」に遷移し、その動作が終了したことに対応して発せられる内部イベント「動作終了」に対応して標準の状態に遷移することなどが記述されている。

## 【0019】

図 1 に戻り、ロボット思考系制御部 11 は、図 2 に示すように、思考系処理部 41 と感情モデル 42 から構成されており、当該ロボットの感情を管理する。思考系処理部 41 は、運動系処理部 31 から入力される動作イベントや各種センサ 1 の検出結果、自己の内部で生成する内部イベント（一定の時間毎、周期的に発生させるイベント等）を感情モデル 42 に照らし合わせてロボットの感情を遷移させ、現在の感情を感情ステートとして音声合成コントロール情報選択部 12 に出力する。思考系処理部 41 はまた、各種センサ 1 の検出結果などに対応して、運動系処理 31 に動作に関する指令を出力する。思考系処理部 41 はさらに、例えば、動作イベント「頭を叩かれた」に対応して音声合成対象文「痛い」を生成するように、動作イベントや各種センサ 1 の検出結果などに対応し、当該ロボットが発声する音声合成対象文を生成して言語処理部 14 に出力する。

## 【0020】

感情モデル 42 には、図 4 に示すように、標準の状態から各種の感情に遷移するときの条件が記述されている。例えば、標準状態において、動作イベント「頭を叩かれた」が発生した場合、感情「怒り」に遷移することや、動作イベント「頭を撫でられた」が発生した場合、感情「嬉しい」に遷移し、その後、所定の時間以上動作イベントが発生されないことに対応して発せられる内部イベントに対応して標準の状態に遷移することなどが記述されている。

## 【0021】

図 1 に戻り、音声合成コントロール情報選択部 12 は、ロボット運動系制御部 10 から入力される動作ステート、および、ロボット思考系制御部 11 から入力される感情ステートに基づいて、音声合成コントロール情報テーブル 13 に多数用意されているフィールドの中から、最も適切な音声合成コントロール情報が格納されているフィールドを選択する。なお、この選択に際しては、動作ステートおよび感情ステート以外のパラメータを追加して、選択するフィールドを決定するようにしてもよい（詳細は、図 8 乃至図 10 を参照して後述する）。

## 【0022】

音声合成コントロール情報テーブル 13 には、動作ステートおよび感情ステート、その他のパラメータ（後述）のあらゆる組み合わせに対応した多数のフィー

ルドが設けられており、音声合成コントロール情報選択部 1 2 に選択されたフィールドに格納している選択情報を言語処理部 1 4 に出力し、音声合成コントロール情報を規則音声合成部 1 5 に出力する。

#### 【 0 0 2 3 】

フィールドには、図 5 に示すように、大別して選択情報と音声合成コントロール情報が格納されている。選択情報は、単語マッピング辞書 ID と文体変換ルール ID からなる。音声合成コントロール情報は、素片データ ID、音節セット ID、声の高さパラメータ、アクセントの強さパラメータ、声立ての強さパラメータ、および、発話スピードパラメータからなる。

#### 【 0 0 2 4 】

単語マッピング辞書 ID は、単語マッピング辞書データベース 5 4（図 6）に予め用意されている、幼児語の単語マッピング辞書、大阪弁の単語マッピング辞書、女子高生言葉の単語マッピング辞書、猫の単語マッピング辞書などの複数の辞書のなかで、単語変換部 5 3（図 6）で使用される辞書を特定するための情報である。単語マッピング辞書は、後述するロボットの個性情報などに対応して切り替えられるものであり、標準語で表記されている音声合成対象文に含まれる単語を、別の単語に置換するための辞書である。例えば、幼児語の単語マッピング辞書は、音声合成対象文に含まれる単語「車」を「ブーブ」に置換する。

#### 【 0 0 2 5 】

文体変換ルール ID とは、文体変換ルールデータベース 5 6（図 6）に予め用意されている、女性言葉への変換ルール、男性言葉への変換ルール、幼児語への変換ルール、大阪弁への変換ルール、女子高生言葉への変換ルール、猫語への変換ルールなどの複数のルールなかで、文体変換部 5 5（図 6）で使用されるルールを特定するための情報である。文体変換ルールは、後述するロボットの個性情報などに対応して切り替えられるものであり、音声合成対象文に含まれる文字列を、別の文字列に変換するためのルールである。例えば、猫語への文体変換ルールは、音声合成対象文中の語尾「\*\*です」を「\*\*ニャ」に変換する。

#### 【 0 0 2 6 】

音声合成コントロール情報に含まれる素片データ ID は、規則音声合成部 1 5

で使用される音声素片を特定するための情報である。音声素片としては、女性の声、男性の声、子供の声、囁れている声、機械的な声、その他の声のデータが、規則音声合成部 15 の内部に予め用意されている。

#### 【0027】

音節セット ID は、規則音声合成部 15 で使用される音節セットを特定するための情報である。音節セットとしては、例えば、基本 266 音節セットと簡略 180 音節セットが用意されている。簡略 180 音節セットでは、基本 266 音節セットよりも発声可能な音韻の数が制限されており、例えば、言語処理部 14 から入力される音声合成対象文「リンゴ」を「にんご」と発声する。このように、発声する音韻を制限することにより、幼児のような舌足らずな発声を表現することができる。

#### 【0028】

声の高さパラメータは、規則音声合成部 15 で合成される音声のピッチ周波数を指定するための情報である。アクセントの強さパラメータは、規則音声合成部 15 で合成される音声のアクセントの強度を指定するための情報である。この値が大きければ、アクセントの起伏が激しい発声が行われ、値が小さければ、平坦なアクセントで発声が行われる。

#### 【0029】

声立ての強さパラメータは、規則音声合成部 15 で合成される音声の声立ての強度を指定するための情報である。この値が大きければ、声立てが多くなり、値が小さければ、声立てが少なくなる。発話スピードパラメータは、規則音声合成部 15 で合成される音声の発話スピードを指定するための情報である。

#### 【0030】

図 1 に戻り、言語処理部 14 は、ロボット思考系制御部 11 から入力される音声合成対象文を文法的に解析した後、音声合成コントロール情報に基づいて所定の箇所を変換し、規則音声合成部 15 に出力する。

#### 【0031】

図 6 は、言語処理部 14 の構成例を示している。ロボット思考系制御部 11 からの音声合成対象文は、文体解析部 51 に入力される。また、音声合成コントロ

ール情報テーブル 1 3 からの選択情報は単語変換部 5 3 および文体変換部 5 5 に入力される。文体解析部 5 1 は、音声合成対象文を解析用辞書 5 2 を用いて形態素解析し、単語変換部 5 3 に出力する。解析用辞書 5 2 には、規則音声合成に必要なとなる情報、例えば、単語（形態素）の読み、アクセント型、および品詞などの他、各単語の固有の単語 I D が記述されている。

## 【 0 0 3 2 】

単語変換部 5 3 は、選択情報に含まれる単語マッピング辞書 I D に対応する辞書を単語マッピング辞書データベース 5 4 から読み出し、文体解析部 5 1 からの形態素解析された音声合成対象文に含まれる単語のうち、読み出した単語マッピング辞書に変換する単語が記述されているものだけを置換して文体変換部 5 5 に出力する。

## 【 0 0 3 3 】

文体変換部 5 5 は、選択情報に含まれる文体変換ルール I D に対応するルールを文体変換ルールデータベース 5 6 から読み出し、読み出した文体変換ルールに従って、単語変換部 5 3 からの単語変換が施された音声合成対象文を変換し、規則音声合成部 1 5 に出力する。

## 【 0 0 3 4 】

図 1 に戻り、規則音声合成部 1 5 は、音声合成コントロール情報テーブル 1 3 から入力される音声合成コントロール情報に従って、言語処理部 1 4 から入力される音声合成対象文に対応する音声信号を合成する。合成された音声信号は、スピーカ 1 6 から放音される。

## 【 0 0 3 5 】

制御部 1 7 は、ドライブ 1 8 を制御して、磁気ディスク 1 9、光ディスク 2 0、光磁気ディスク 2 1、または半導体メモリ 2 2 に記憶されている制御用プログラムを読み出し、読み出した制御用プログラムに基づいて各部を制御する。

## 【 0 0 3 6 】

次に、本発明を適用したロボットの処理について、図 7 のフローチャートを参照して説明する。この処理は、例えば、ユーザに頭を叩かれたことを、各種センサ 1 の感圧センサ 5 が検出し、検出結果がロボット運動系制御部 1 0 の運動系処

理部 31 に入力されたときに開始される。

【0037】

ステップ S1 において、運動系処理部 31 は、感圧センサ 5 の検知結果が所定の閾値以上の力であることに対応して、動作イベント「頭を叩かれた」の発生を判定し、ロボット思考系制御部 11 の思考系処理部 41 に通知する。運動系処理部 31 はまた、動作イベント「頭を叩かれた」を動作モデル 32 に照らし合わせてロボットの動作「起き上がる」を決定し、動作ステートとして音声合成コントロール情報選択部 12 に出力する。

【0038】

ステップ S2 において、ロボット思考系制御部 11 の思考系処理部 41 は、運動系処理部 31 から入力された動作イベント「頭を叩かれた」を感情モデル 42 に照らし合わせて、感情を「怒り」に遷移させ、現在の感情を感情ステートとして音声合成コントロール情報選択部 12 に出力する。思考系処理部 41 はまた、動作イベント「頭を叩かれた」に対応して音声合成対象文「痛い」を発生して言語処理部 14 の文体解析部 51 に出力する。

【0039】

ステップ S3 において、音声合成コントロール情報選択部 12 は、運動系処理部 31 から入力された動作ステート、および、思考系処理部 41 から入力された感情ステートに基づいて、音声合成コントロール情報テーブル 13 に多数用意されているフィールドの中から、最も適切な音声合成コントロール情報が格納されているフィールドを選択する。音声合成コントロール情報テーブル 13 は、選択されたフィールドに格納している選択情報を音声処理部 14 に出力し、音声合成コントロール情報を規則音声合成部 15 に出力する。

【0040】

ステップ S4 において、言語処理部 14 の文体解析部 51 は、音声合成対象文を解析用辞書 52 を用いて形態素解析し、単語変換部 53 に出力する。ステップ S5 において、単語変換部 53 は、選択情報に含まれる単語マッピング辞書 ID に対応する辞書を単語マッピング辞書データベース 54 から読み出し、文体解析部 51 からの形態素解析された音声合成対象文に含まれる単語のうち、読み出し

た単語マッピング辞書に変換する単語が記述されているものだけを置換して文体変換部 55 に出力する。ステップ S6 において、文体変換部 55 は、選択情報に含まれる文体変換ルール ID に対応するルールを文体変換ルールデータベース 56 から読み出し、読み出した文体変換ルールに従って、単語変換部 53 からの単語変換が施された音声合成対象文を変換し、規則音声合成部 15 に出力する。

## 【0041】

ステップ S7 において、規則音声合成部 15 は、音声合成コントロール情報テーブル 13 から入力された音声合成コントロール情報に従って、言語処理部 14 から入力された音声合成対象文に対応する音声信号を合成し、スピーカ 16 から放音させる。

## 【0042】

以上のような処理により、ロボットはあたかも感情を有するかのように動作し、また、感情の変化や動作に応じて話し方を変化するようになる。

## 【0043】

次に、音声合成コントロール情報選択部 12 の選択処理に際し、動作ステートおよび感情ステート以外のパラメータを追加する方法について、図 8 乃至図 10 を参照して説明する。

## 【0044】

図 8 は、図 1 の構成例に対して、ロボットに個性を設定するための構成として、通信ポート 61 乃至個体情報メモリ 63 を追加した構成例を示している。通信ポート 61 は、外部装置（パーソナルコンピュータなど）と個性情報を通信するインタフェースであり、例えば、RS-232C、USB、IEEE1394等の通信規格に適合したものが考えられる。通信制御部 62 は、通信ポート 61 を介した外部との情報通信を所定のプロトコルに従って制御し、受信する個性情報をロボット思考系制御部 11 に出力する。個性情報メモリ 13 は、フラッシュメモリなどの書き換え可能な不揮発性メモリを用い、記憶した個性情報を音声合成コントロール情報選択部 12 に出力する。

## 【0045】

外部から供給する個性情報としては、例えば、次のような項目が考えられる。

種別 犬／猫

性別 雄／雌

年齢 子供／大人

気性 激しい／おとなしい

体格 痩せている／太っている

【 0 0 4 6 】

これらの各項目は、0 または 1 に 2 値化されて個性情報メモリ 6 3 に記憶される。また、各項目を 2 値化するのではなく、さらに細分化して設定するようにしてもよい。

【 0 0 4 7 】

個性情報がむやみに書き換えられることを抑止するために、書き換えの回数を制限するようにしてもよいし、書き換えに際してパスワードを設けるようにしてもよい。また、通信ポート 6 1 および通信制御部 6 2 を設けずに、個性情報を予め書き込んだROMから成る個性情報メモリ 6 3 を製造時に組み込むようにしてもよい。

【 0 0 4 8 】

このような構成とすることにより、設定された個性に応じて、個体毎に異なる音声を出力するロボットを実現することが可能となる。

【 0 0 4 9 】

図 9 は、図 1 の構成例に対して、タイマ 7 1 を追加した構成例を示している。タイマ 7 1 は、ロボットが最初に起動されたとき以降の経過時間をカウントして音声合成コントロール情報選択部 1 2 に出力する。なお、タイマ 7 1 は、最初に起動されたとき以降の稼働中の時間だけのカウントするようにしてもよい。

【 0 0 5 0 】

このような構成とすることにより、出力する音声を経年変化させるロボットを実現することが可能となる。

【 0 0 5 1 】

図 1 0 は、図 1 の構成例に対して、経験値計算部 8 1 および経験値メモリ 8 2 を追加した構成例を示している。経験値計算部 8 1 は、思考系処理部 4 1 が感情



を標準から他の状態に遷移させる毎、遷移した感情ステート毎に遷移回数をカウントアップして経験値メモリ 8 2 に記録する。例えば、図 4 に示した感情モデル 4 2 のように感情ステートが 4 種類で有れば、4 項目の遷移回数が経験値メモリ 8 2 に記録される。音声合成コントロール情報選択部 1 2 には、各感情ステートへの遷移回数を通知するようにしてもよいし、最も遷移回数の多い感情ステートを通知するようにしてもよい。

## 【0052】

このような構成とすることにより、例えば、頻繁に叩かれていて、感情ステート「怒り」への遷移回数が多いいロボットは、怒りっぽい話し方をするようにすることができ、頻繁に撫でられていて、感情ステート「嬉しい」への遷移回数が多いいロボットは、楽しそうな話し方をするようにすることができる。

## 【0053】

なお、図 8 乃至図 1 0 に示した構成例は、適宜組み合わせることが可能である。

## 【0054】

また、各種センサ 1 の検出結果をパラメータ化して、音声合成コントロール情報選択部 1 2 に供給することにより、外部の状況に応じて話し方を変化させるようにしてもよい。例えば、外気温センサ 2 が検出した外気温が所定の温度以下である場合、寒そうな震えた声を発声するようにしてもよい。

## 【0055】

さらに、各種センサ 1 の検出結果をパラメータ化して、履歴として記録し、その履歴を音声合成コントロール情報選択部 1 2 に供給するようにしてもよい。これにより、例えば、外気温が所定の温度以下である履歴が多いロボットは、東北弁を話すようになるようにしてもよい。

## 【0056】

ところで、上述した一連の処理は、ハードウェアにより実行させることもできるが、ソフトウェアにより実行させることもできる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをイ

インストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

【 0 0 5 7 】

この記録媒体は、図 1 に示すように、コンピュータとは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク 1 9 (フロッピーディスクを含む)、光ディスク 2 0 (CD-ROM(Compact Disc-Read Only Memory)、DVD(Digital Versatile Disc)を含む)、光磁気ディスク 2 1 (MD(Mini Disc)を含む)、もしくは半導体メモリ 2 2 などよりなるパッケージメディアにより構成されるだけでなく、コンピュータに予め組み込まれた状態でユーザに提供される、プログラムが記録されているROMやハードディスクなどで構成される。

【 0 0 5 8 】

なお、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に従って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

【 0 0 5 9 】

【発明の効果】

以上のように、本発明の音声合成装置、音声合成方法、および記録媒体のプログラムによれば、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択し、選択した制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成するようにしたので、感情や動作等に対応して話し方を変化させることができ、よりリアルに生命体を模擬したロボットを実現することが可能となる。

【図面の簡単な説明】

【図 1】

本発明を適用したロボットの音声合成に関わる部分の構成例を示すブロック図である。

【図 2】

図 1 のロボット運動系制御部 1 0 およびロボット思考系制御部 1 1 の構成例を示すブロック図である。

【図 3】

図 2 の動作モデル 3 2 を説明するための図である。

【図 4】

図 2 の感情モデル 4 2 を説明するための図である。

【図 5】

音声合成コントロール情報を説明するための図である。

【図 6】

言語処理部 1 4 の詳細な構成例を示すブロック図である。

【図 7】

本発明を適用したロボットの動作を説明するフローチャートである。

【図 8】

本発明を適用したロボットの音声合成に関わる部分の他の構成例を示すブロック図である。

【図 9】

本発明を適用したロボットの音声合成に関わる部分のさらに他の構成例を示すブロック図である。

【図 1 0】

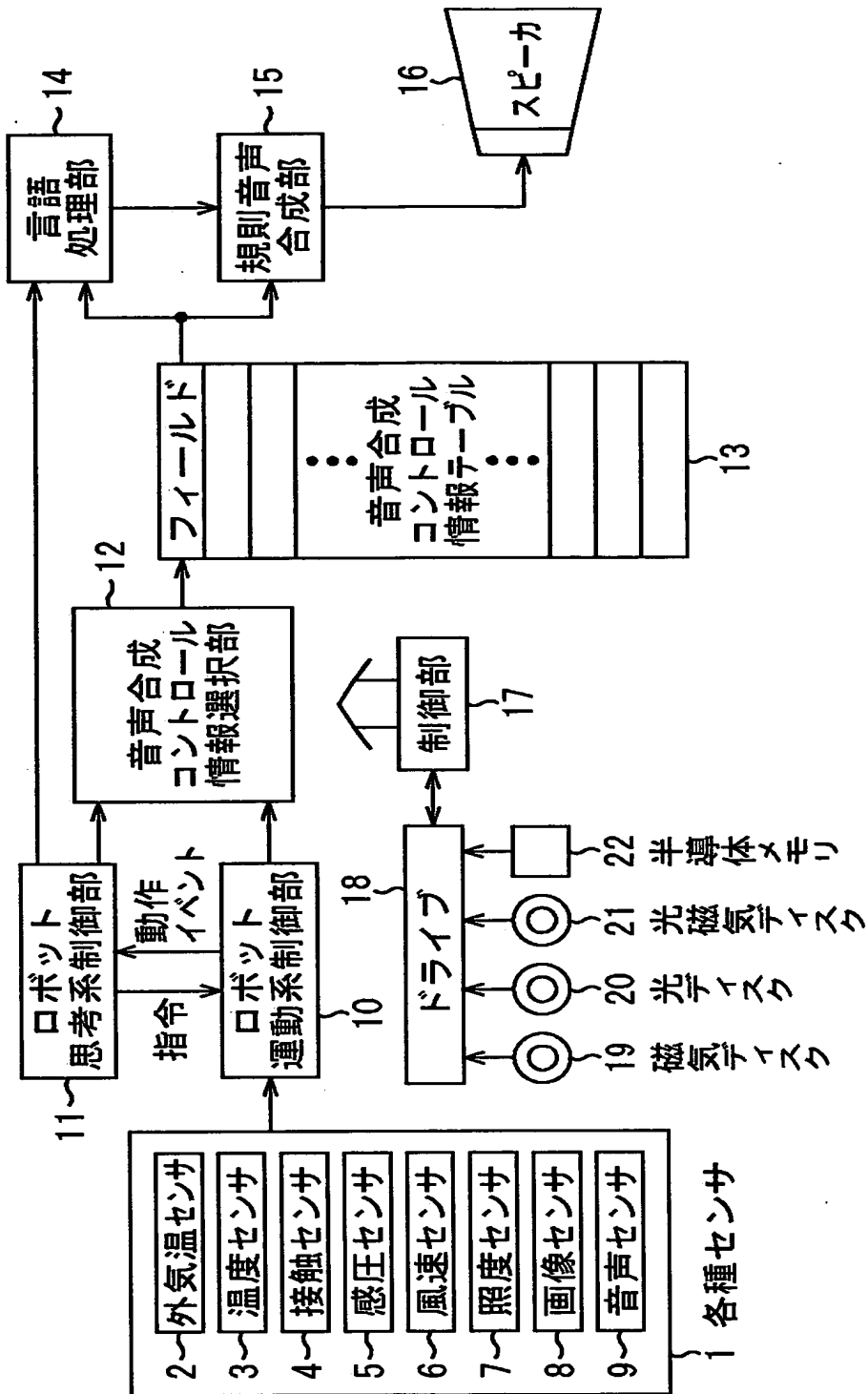
本発明を適用したロボットの音声合成に関わる部分のさらに他の構成例を示すブロック図である。

【符号の説明】

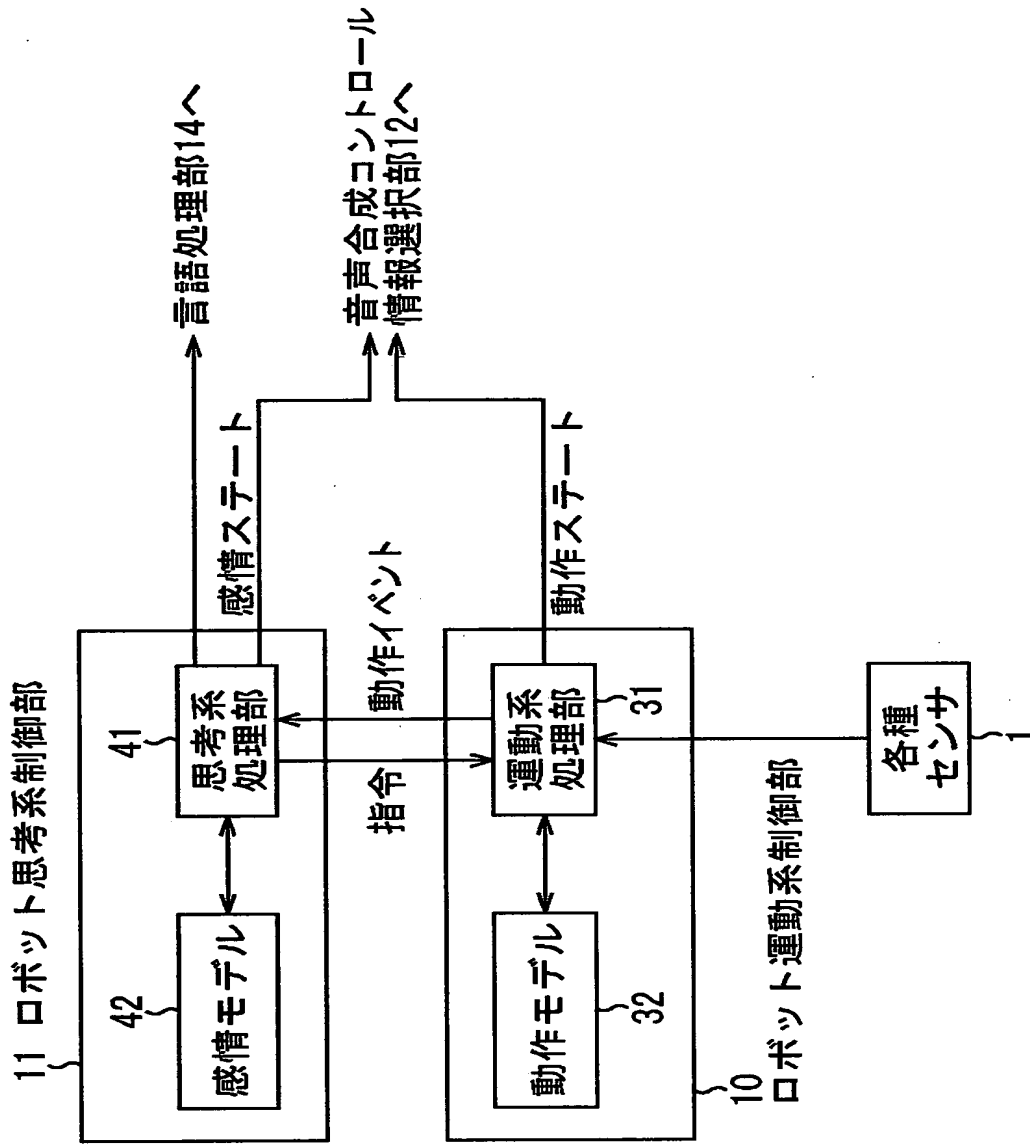
1 各種センサ, 1 0 ロボット運動系制御部, 1 1 ロボット思考系制御部, 1 2 音声合成コントロール情報選択部, 1 3 音声合成コントロール情報テーブル, 1 4 言語処理部, 1 5 規則音声合成部, 1 7 制御部, 3 1 運動系処理部, 3 2 動作モデル, 4 1 思考系処理部, 4 2 感情モデル, 5 1 文体解析部, 5 3 単語変換部, 5 5 文体変換部, 6 1 通信ポート, 6 2 通信制御部, 6 3 個体情報メモリ, 7 1 タイマ, 8 1 経験値計算部, 8 2 経験値メモリ

【書類名】 図面

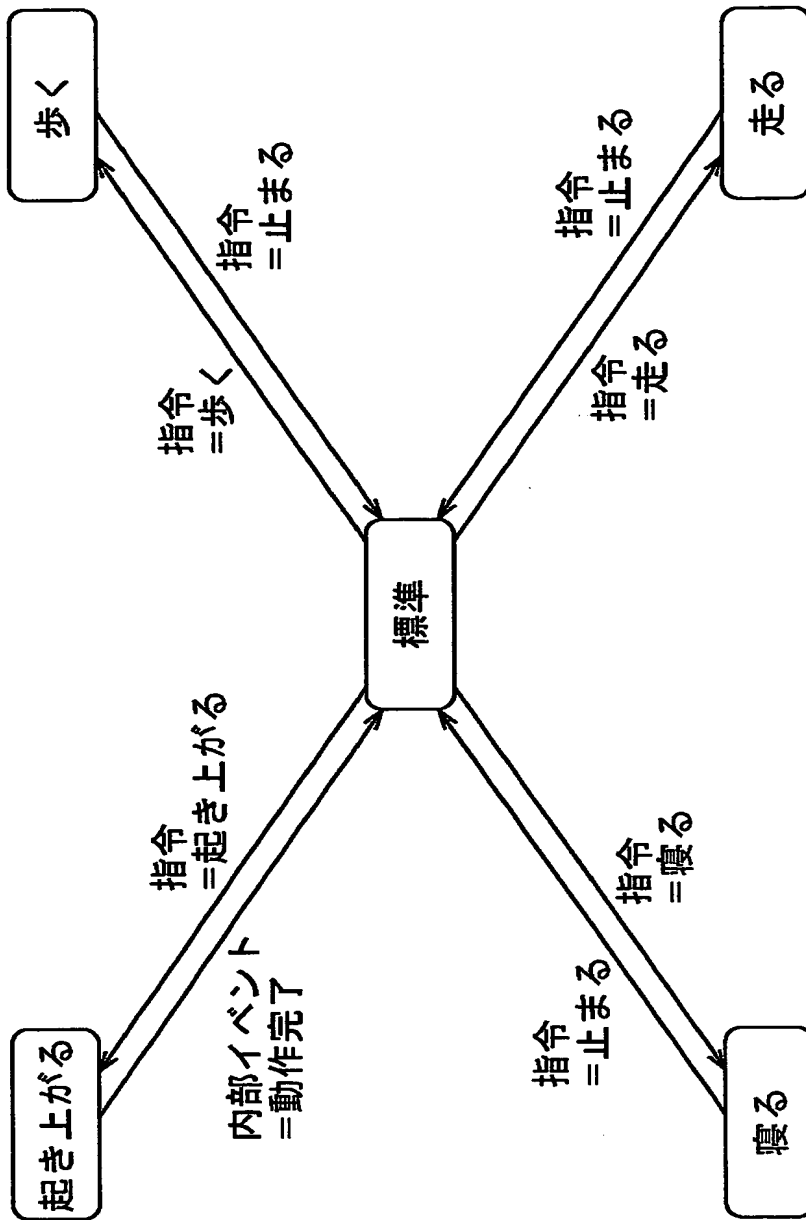
【図 1】



【図 2】

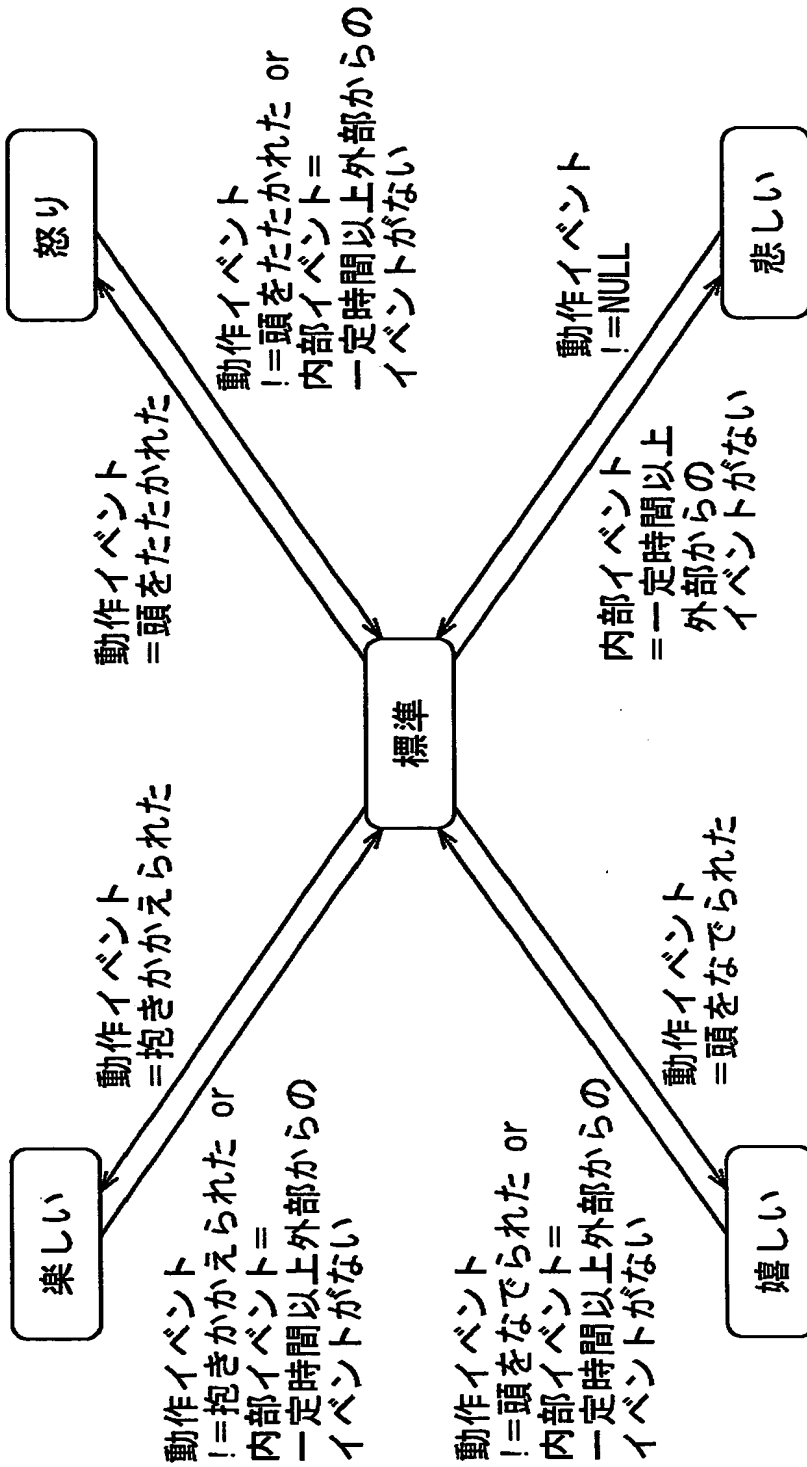


【図 3】



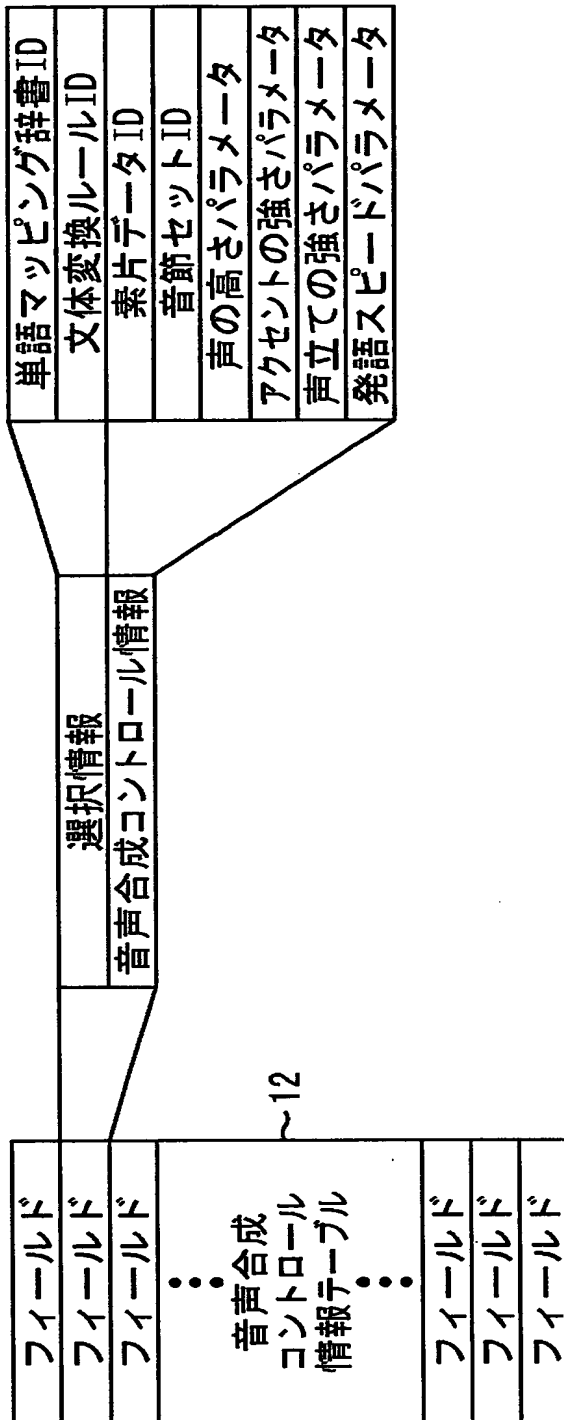
動作モデル 32

【図 4】



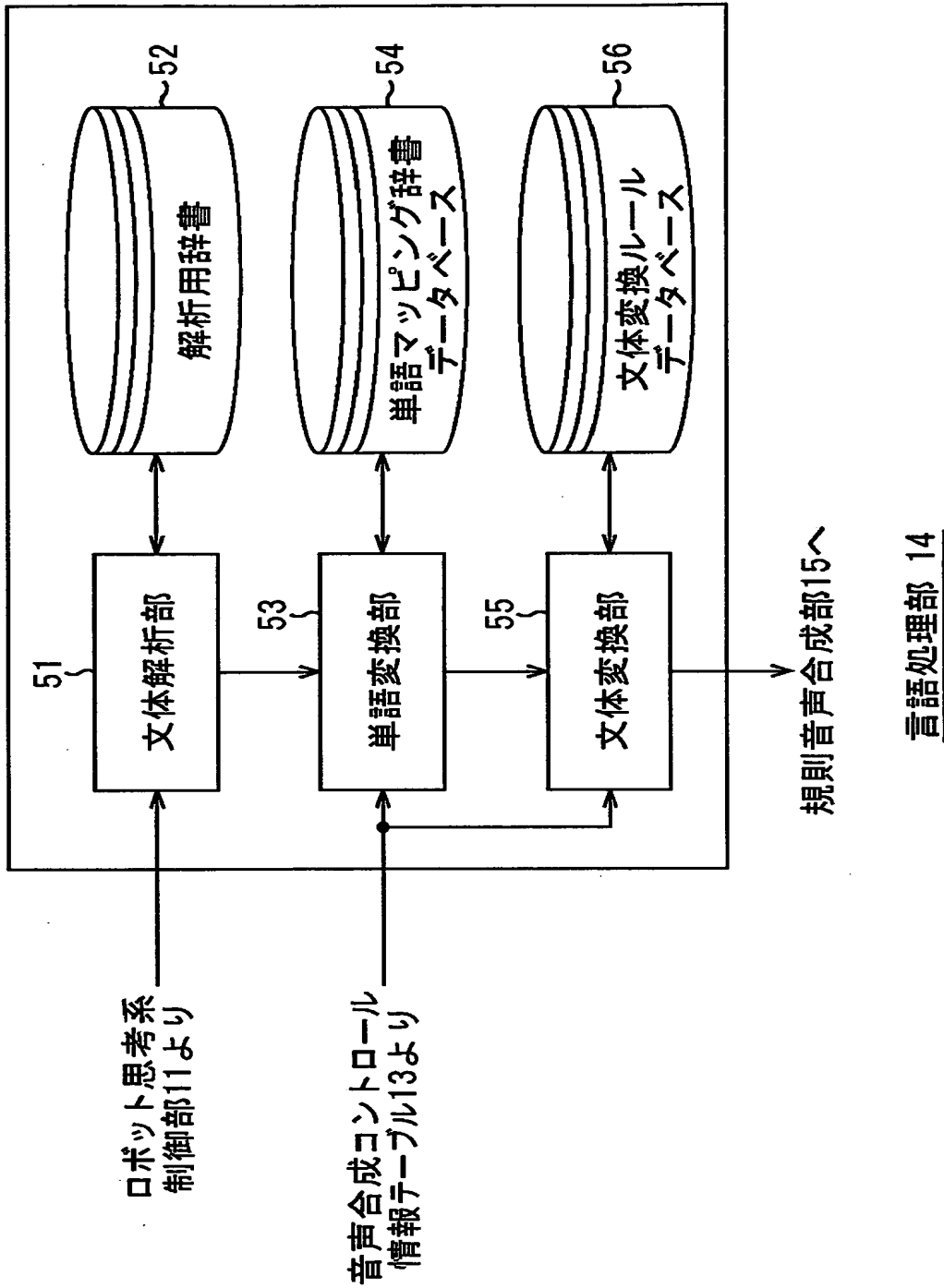
感情モデル 42

【図 5】

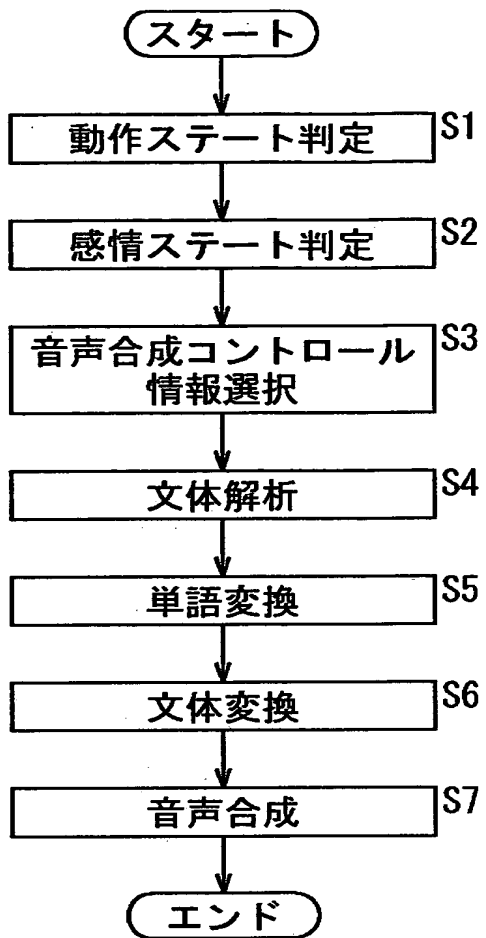




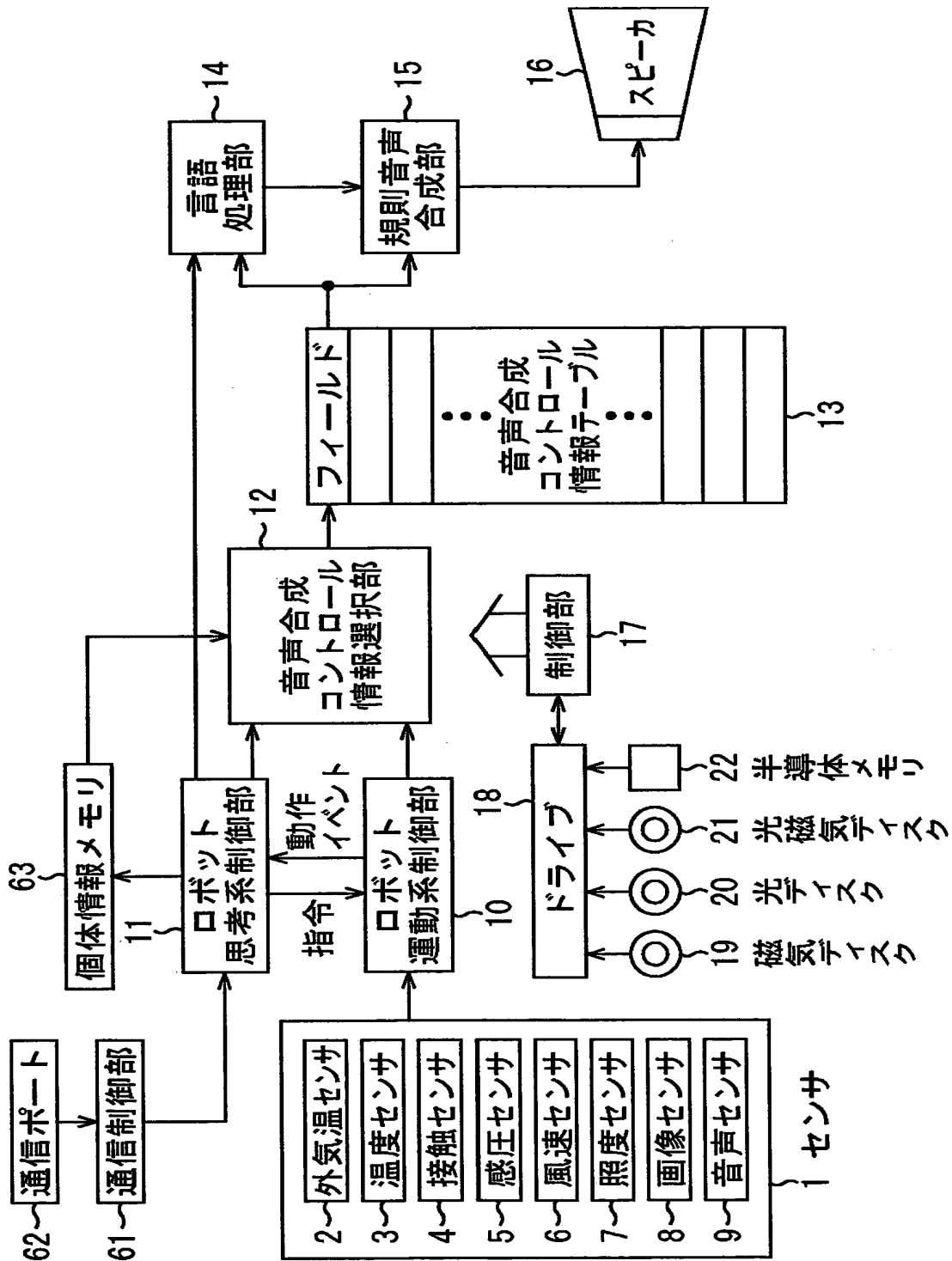
【図 6】



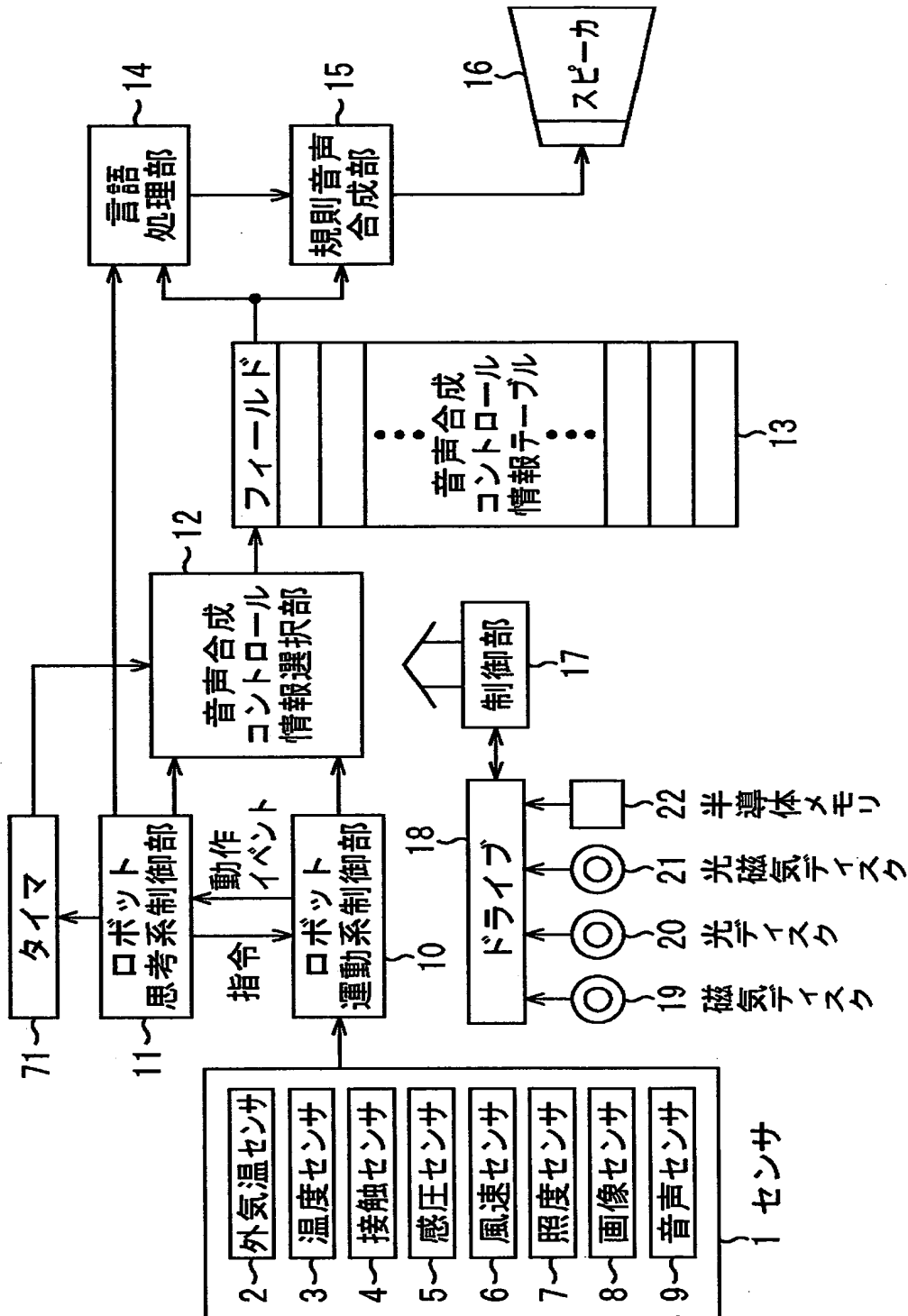
【図 7】



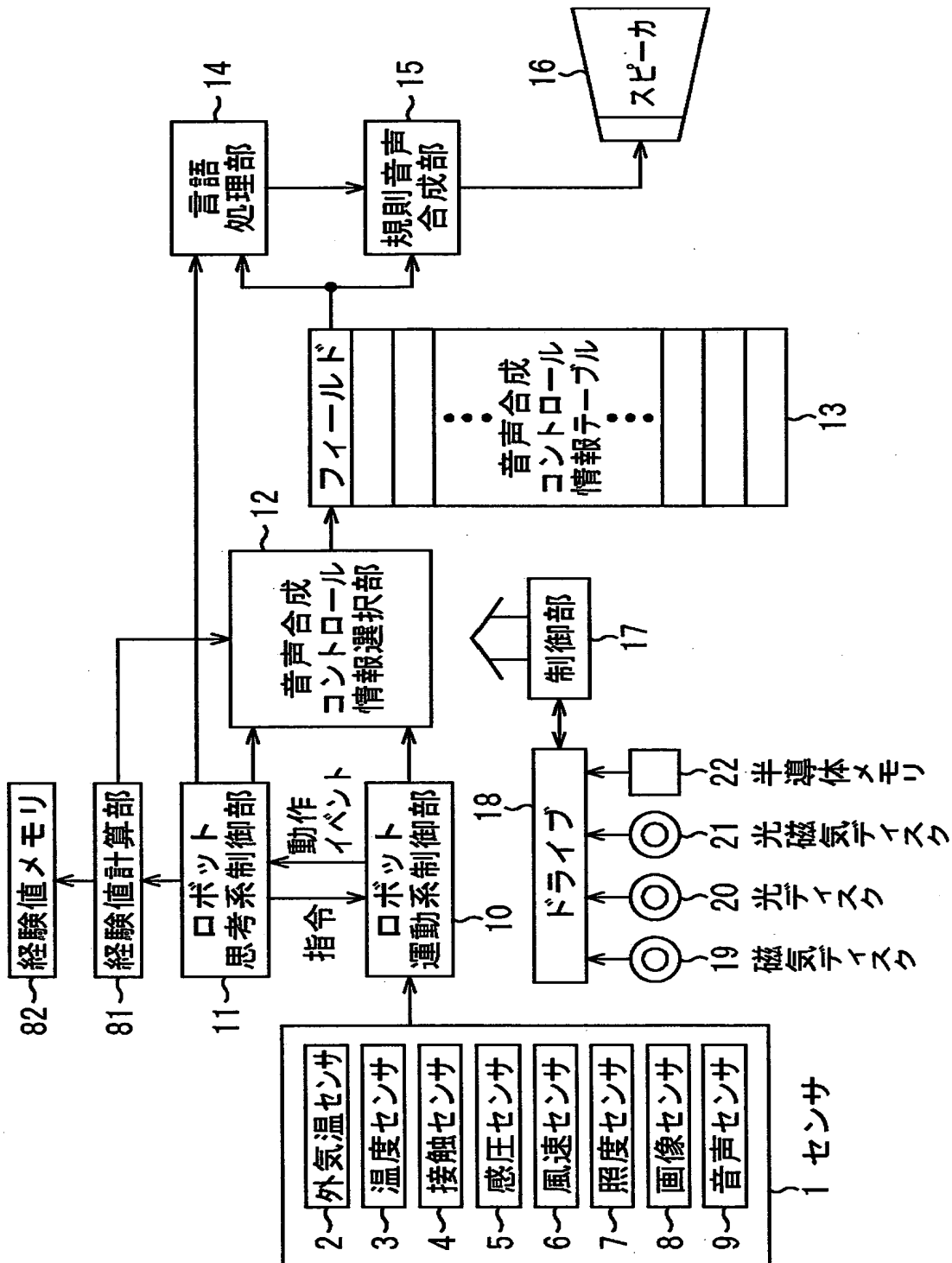
【図 8】



【図 9】



【図 1 0】



【書類名】 要約書

【要約】

【課題】 感情や動作等に対応して話し方を変化するロボットを実現する。

【解決手段】 各種センサ 1 は、外部の状況やロボットに対する作用を検出して、検出結果をロボット運動系制御部 1 0 に出力する。ロボット運動系制御部 1 0 は、動作モデル 3 2 に基づき、動作ステートを決定する。ロボット思考系制御部 1 1 は、感情モデルに基づき、感情ステートを決定する。音声合成コントロール情報選択部 1 2 は、動作ステート、および感情ステートに基づいて、音声合成コントロール情報テーブル 1 3 上のフィールドを決定する。言語処理部 1 4 は、ロボット思考系制御部 1 1 からの音声合成対象文を文法的に解析した後、音声合成コントロール情報に基づいて所定の箇所を変換し、規則音声合成部 1 5 に出力する。規則音声合成部 1 5 は、音声合成対象文に対応する音声信号を合成する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 2 1 8 5]

1. 変更年月日	1 9 9 0 年 8 月 3 0 日
[変更理由]	新規登録
住 所	東京都品川区北品川 6 丁目 7 番 3 5 号
氏 名	ソニー株式会社